US-PAT-NO: **6184926**

DOCUMENT-IDENTIFIER: US 6184926 B1

TITLE: System and method
for detecting a human face in
                        uncontrolled
environments

---------- KWIC ----------

US Patent No. - PN (1):
  6184926

Detailed Description Text - DETX (5):
   The video signal 115 is input to the
frame grabber 120. In one embodiment,
the frame grabber 120 may comprise a Meteor
Color Frame Grabber, available from
Matrox. The frame grabber 120 operates to
convert the analog video signal 115
into a digital image stored within the
memory 135, which can be processed by
the video processor 140. For example, in
one implementation, the frame grabber
120 may convert the video signal 115 into a
640.times.480 (NTSC) or
768.times.576 (PAL) color image. The color
image may consist of three color
planes, commonly referred to as YUV or YIQ.

Each **pixel** in a color plane may
have 8 bits of resolution, which is
sufficient for most purposes. Of course, a
variety of other digital image formats and
resolutions may be used as well, as
will be recognized by one of ordinary
skill.

Detailed Description Text - DETX (17):
   To extract moving objects within the
video image stored in memory 135, the
background may be modeled as a texture with
the intensity of each point modeled
by a Gaussian distribution with mean .mu.
and variance .sigma., N.sub.b
(.mu.,.sigma.) (step 216). The **pixels** in
the image are classified as
foreground if p(O(x,y).vertline.N.sub.b
(.mu.,.sigma.)).ltoreq.T and as
background if p(O(x,y).vertline.N.sub.b
(.mu.,.sigma.))>T. The observation
O(x,y) represents the intensity of the
pixels at location (x,y), and T is a
constant (step 212).

Detailed Description Text - DETX (18):
   The connectivity analysis (step 213) of
the "foreground" **pixels** generates
connected sets of **pixels,** i.e. sets of
pixels that are adjacent or touching.
Each of the above sets of **pixels** describe a
foreground region. Small
foreground regions are assumed to be due to

shadow, camera noise and lighting
variations and are removed.

Detailed Description Text - DETX (22):
   where the observations O(x,y) are the
pixel intensities at coordinates (x,y)
in the foreground regions and
P(O(x,y).vertline..lambda..sub.i) is the
likelihood functions for the i.sup.th model
301.

Detailed Description Text - DETX (28):
   For each set of parameters
(x.sub.0,y.sub.0,a,b) a rectangular
template (W
in FIG. 10) is defined by the set of
parameters
(x.sub.0,y.sub.0,.alpha.a,.alpha.b), where
x.sub.0 and y.sub.0 are the
coordinates of the center of the rectangle
and .alpha.a,.alpha.b are the width
and length of the rectangle, and .alpha.
is some constant (see FIG. 10). In
each area bounded by x.sub.j-1,x.sub.j,
R.sub.out,j is the set of **pixels**
outside the ellipse template and inside the
rectangle template and R.sub.in,j
is the set of **pixels** inside the ellipse
template (FIG. 10). The regions
R.sub.in,j and R.sub.out,j locally classify
the image in "face" and "non face"
regions. Based on the above discussion,
the likelihood function

P(O(x,y).vertline..lambda..sub.i) for the
model .lambda..sub.i is determined by
the ratio of the number of foreground
pixels classified as "face" and
background **pixels** classified as "non face"
in each area bounded by
x.sub.j-1,x.sub.j, (where j=1 to i), over
the total number of **pixels** in "face"
and "non face" regions (step 403). This is
described in Equation (2) below:
##EQU2##

Detailed Description Text - DETX (66):
    Stage 202 of FIG. 2 illustrates the
steps of the eye detection approach that
may be used according to the present
invention. In step 221, the **pixel**
intensities inside the face regions are
compared to a threshold .theta., and
pixels with intensities lower than .theta.
are extracted from the face region.
In step 222, and as shown in FIG. 7A, the
connectivity analysis of the
extracted **pixels** generates connected sets
of **pixels** (e.g., **pixels** 701), i.e.
sets of **pixels** that are adjacent or
touching. Each of these connected sets of
pixels 701 describe a low intensity region
of the face.

Detailed Description Text - DETX (67):
    In step 223, the **pixel** regions 701
resulting from steps 221-222 are filtered

with respect to the region size. Regions
having a small number of **pixels** due
to camera noise or shadows are removed.
Large regions generally cannot
represent eyes, but instead correspond in
general to hair. The size of the
regions selected at this stage is in the
interval [.theta..sub.m,.theta..sub.M
] where .theta..sub.m is the minimum and
.theta..sub.M is the maximum number of
pixels allowed by our system to describe a
valid eye region. Threshold values
.theta..sub.m,.theta..sub.M are determined
based on the size of the ellipse
that characterizes the head region (the
ellipse being generated iteratively in
step 215). The end result of step 223 is
an image 702, such as that shown in
FIG. 7B.